ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Physics Procedia 3 (2010) 1781–1789

Physics Procedia

www.elsevier.com/locate/procedia

Structure and External Factors of Chinese City Airline Network

Hong-Kun Liu [1], Xiao-Li Zhang [2], Tao Zhou [3 4*]

[1] School of Statistics, South Western University of Finance and Economics, Chengdu 610074, China
[2] College of Economics & Management，Shanghai Ocean University, Shanghai 200090, China
[3] Web Sciences Center，University of Electronic Science and Technology of China, Chengdu 610054, China
[4] Department of Modern Physics, University of Science and Technology of China, Hefei 230026, China

Abstract

We investigate the structural properties of Chinese city airline network (CCAN), where nodes and edges denote cities and direct flights. The degree distribution follows a double power law and a clear hierarchical layout is observed. The population exhibits a weakly positive correlation with the number of flights, yet it does not show obvious correlation with the transportation flow. The distance is an important parameter in CCAN, that is, the number of flights decays fast with the increasing of the distance. In comparison, the tertiary industry has the most important influence on the Chinese air passenger transportation. Statistically speaking, when the tertiary industry value increases by 1%, the next period's volume will increase by 0.73%.

Keywords: complex networks; airline network; structural properties; external factors

1. Intrduction

Being the basis of transportation industries in 20th century, the air transportation has been one of the most successful industries for its great achievements in the techniques and services. The advantages of air transportation, compared with others, are threefold: breadth, depth and speed. Since the first scheduled flights from Tampa to St. Petersburg in the United States in 1914, the aircrafts and the engines have been constantly improved and refined to enhance the carrying capacity, range and speed. They have also promoted the formation of the world airline network [1]. Air transportation is not only a mode of transport, but also high-quality resources. Utilizing air transportation, regional economic could transcend the geographical boundaries of space, and directly participate in the international division of labor and international economic cycle. It attracts many elements of productivity in the world. Therefore, studying the structures of airline networks, especially external factors that affect the evolution and organization of the networks, is of practical significance.

The study on airline networks is of common interests of Economics [2-4], Geography [5-8], Transportation [9-11], Operations [12-15] and Physics [16-18]. The researchers study the characteristics, design, advantages and

* *E-mail address*: zhutou@ustc.edu.

doi:10.1016/j.phpro.2010.07.019

disadvantages of airline network structures from different viewpoints, as well as economic and social impacts arising from it [19]. In contrast, analysis about the effects of external factors on the airline network structure is rarely reported. Based on the topology, this paper will analyze how population, economic size and distance between two cities affect the connections of Chinese city airline network (CCAN).

2. Structural Properties

CCAN consists of 121 nodes and 1378 edges, where nodes and edges represent cities and airlines [20]. Our data is collected from the 8 largest air companies in 2006. The average degree is 11.388, and the five largest-degree cities are Beijing, Shanghai, Guangzhou, Shenzhen and Chengdu. We define the cities with degree less than the average as small-degree cities, while the others as large-degree ones. The average path length and clustering coefficient are $L = 2.263$ and $C = 0.748$, which indicates that CCAN is a typical small-world network. As shown in Fig. 1, the degree distribution follows a double power-law, indicating that a few hub nodes have very large degrees.

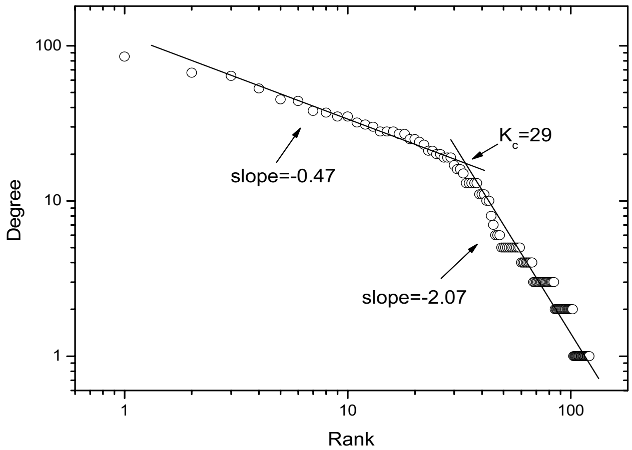

Figure 1: Degree distribution of CCAN in the Zipf's plot [21]. The Y-axis denotes node's degree, and X-axis denotes the ranks of nodes according to their degrees.

Meanwhile, CCAN has a clear hierarchical layout, whose backbone is radially interconnected and inter-provincial airlines. Some provinces have already formed radial regional airline networks where provincial capitals or major cities are the hubs and the local airlines constitute the main body. These local sub-networks connect with the whole national network with trunk lines.

According to the analysis of the data, we get the following results:

(a). Being the four largest-degree cities, Beijing, Shanghai, Guangzhou and Shenzhen are national hubs. Compared with other large-degree cities, the percentages of connections to small-degree cities of these four hubs are much higher.

(b). The remaining 34 large-degree cities are more likely to connect with other large-degree cities, except three special cities: Kunming, Urumqi and Hohhot. Though they connect small-degree cities with high percentages, the majority of the connected small-degree cities are located inside the corresponding provinces. Therefore, they are recognized as regional central cities. Except of Kunming, Urumqi and Hohhot, in CCAN, the city with higher percentage of connections to small-degree cities is more central.

(c). In Yunnan and Sinkiang, small-degree cities do not connect with other cities outside the region, but only connect with the corresponding capital cities and/or interconnect inside the provinces.

(d). Among the 1348 edges, nearly 98% of which have at least one large-degree node as one endpoint. 10 largest-degree cities and 20 highest-capacity flights are shown in Fig. 2. The transportation capacity is defined as the total number of available seats per week provided by air companies between cities. Accordingly, a city's transportation capacity is defined as the total number of available seats of all airlines connected this city.

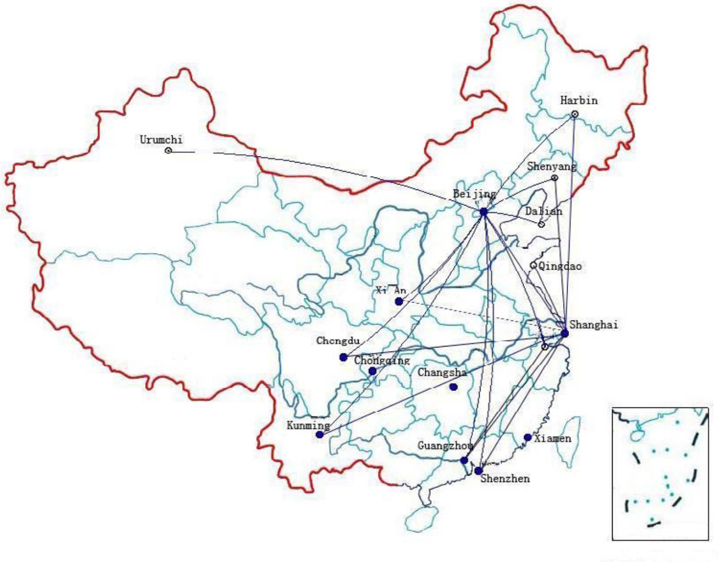

Figure 2: 20 strongest airlines in CCAN

3. Impacts of External Factors

The present researches suggest that only considering network's topology without taking into account external factors is unable to capture the evolution process of airline networks [22, 23]. Accordingly, this section will discuss the external factors and anaysis their impacts on air passenger transportation.

Exchanges of material, energy, human and information among cities occure minute to minute. Interactions between two cities will happen if the exchange of some elements will bring benefit to both of them, meanwhile there are no supply and consumer demands of another city between them. In the air transportation, the requirement of interactions between two cities determines whether we need to establish an airline and strength of interactions reflects the scale of transportation.

There are several interpretations about the regularities in inter-city spatial interaction. Whatever the gravity model, the potential model or the law of retail gravitation, the factors of inter-city spatial interaction include the distance between two cities and their masses which are usually quantified by population or GDP. Here, we concern the airlines between cities. Therefore, external factors affecting airlines are related to population, economic size and distance, and we will discuss the correlations between passenger transportation and population, economic scale and distance.

3.1 Population and airline capacity

In air transportation, the number of passengers is related to the development of commodity economy, traffic location, standard of living, the level of urbanization, and so on. Among all underlying factors, the population is the most foundational one. In general, the larger the city population is, the more likely the city will conncet with other cities. Figure 3 reflects the relationship between urban population and degree. In CCAN, the city's population is basically positive correlated with the airline number. There are exceptions that some cities have more airline numbers but their population are not large. Figure 4 displays the relationship between city's population and its transportation capacity in 2004 (the results are almost the same for 2001-2003). Overall speaking, there is a weakly positive correlation between population and transportation capacity, but there exists some large-capacity cities with small population.

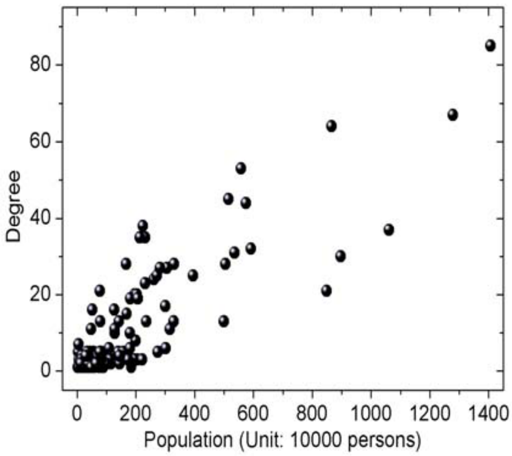

Figure 3: The relationship between the population of a city and its degree.

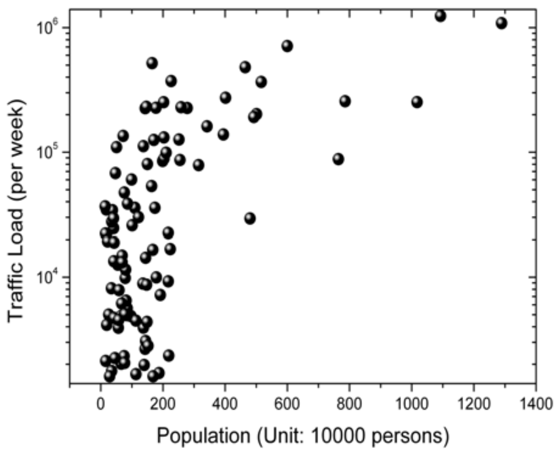

Figure 4: The relationship between the population of a city and the total flow associated with it.

3.2 Distance and airline capacity

Distance is used to describe spatial proximity of two or more things. Moving from one place to another costs both time and energy, and thus many economic and social activities are limited to a certain distance. In the literature, distance plays an important role in transportation networks [24-32]. Transportation network mainly includes road networks [33-36], MTR networks [37], railway networks [38-40], and airline networks. Road networks, MTR networks, railway networks are the different type of transportation networks from airline networks. In these networks, the edges are physical connections, while in airline networks the flights are not physical conncetions. In fact, in railway network, the degree distribution is not power-law because the cost due to distance depresses the long connections. In India, Switzerland and Central Europe, railway networks' degree distributions are exponential [38-40].

Figure 5 is the relationship of airline number and distance between cities. Two cities have at most one airline whatever the number of flights between them. As shown in Fig. 5, the distances assemble in the range from 400km to 1400km. When the distance exceeds 1400km, the airline number decays very fast with the increasing of distance.

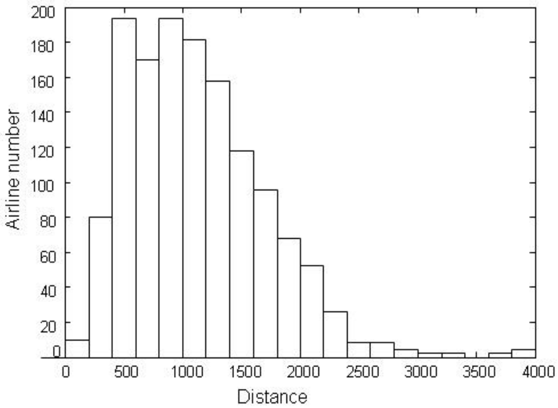

Figure 5: The relation between airline number and distance

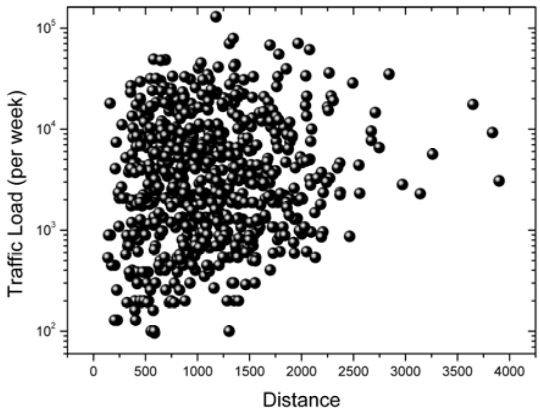

Figure 6: The relation between distance and transportation capacity

With the increasing of distance, the cost of air transportation will increase too. The higher transportation cost is, the greater difficulty of communication between cities is. The result indicated in Fig. 5 is in accordance with the distance-decay principle, which suggests that the interaction strength between the geographical objects is inversely proportional to the distance. In a certain distance range (<2000km), there is no clear correlation between distance and flow. But when the distance exceeds 2000km, transportation capacity will be neither light nor heavy (see Fig. 6). If the number of passengers is small, the long-distance flights can not be well fed, and thus these flights are of high probablility to be given up by companies. On the other hand, the high prices of long-distance flightes drive some passengers to choose other transportation tools, and the long-distance travels are fewer than short-distance ones [41]. Therefore, long-distance airlines usually do not have heavy flows.

3.3 Three industries and transportation capacity

We here decompose economic scale for three industries' values, then respectively discuss the correlation between three industries' values and air passenger volume from 1989 to 2004 in China. Price factor has been excluded in three industries' values (all in 1985 constant prices). Table 1 shows the correlation between Primary, secondary industry values and the air passenger volume.

Table 1 The correlation between Primary, secondary industry values and the air passenger volume

		passenger volume Pearson Correlation	Significance (2-tailed)
primary industry value	Direct Correlation	0.972(**)	.000
	Partial Correlation （Control Variables: tertiary industry volume）	0.137	.626
secondary industry value	Direct Correlation	0.989(**)	.000
	Partial Correlation （Control Variables: tertiary industry volume）	0.322	.261

** Correlation is significant at the 0.01 level (2-tailed).

As shown in Table 1, the correlation coefficient of the primary industry value and air passenger volume is 0.972, and the probability of significance test is 0.000. The correlation coefficient of the secondary industry value and air passenger volume is 0.989, and the significant probability of test is 0.000. These indicate the primary and secondary industry values are all highly correlated with air passenger volume. However, partial correlation analysis excluding the tertiary industry found that the correlation coefficient of the primary industry value and air passenger volume is 0.137, and t test value is 0.626>0.05, suggesting that there is no significant likelihood that the primary industry value is correlated with air passenger volume, namely the primary industry is correlated with air passenger volume through the tertiary industry. Similarly, without considering the tertiary industry, the correlation coefficient of the secondary industrial value and air passenger volume is 0.322, and t test value is 0.261>0.05, also indicating that the secondary industry is correlated with air passenger volume through the tertiary industry. These results suggest that the scale of urban circulation and services have great impacts on scheduling flights.

As a fundamental service industry, the tertiary industry is the general term of various kinds of industries for the social production and people's living by virtue of a certain material and technical equipment services. Being the division of labor and specialization in modern society, each individual will focuses on narrower-scale economic activities. For example, the separation of production and sales causes the whole process to be fulfilled on the basis of transportation, from raw materials procurement, production to consumption. Almost all social and economic activities are based on transportation, not only production and distribution in the primary and secondary industry, but also service lines in the tertiary industry other than the one of transportation. We could imagine that there is correlation between the tertiary industry and transportation. Then, what’s the specific relation between the tertiary industry and the air transportation, being of the rapidest growth in this line.

With the Granger causality test, we will examine the causal relation between the tertiary industry value and air passenger volume, using the relative time-series data from 1989 to 2004 (all in 1985 constant prices). Here, X represents air passenger volume and Y is the tertiary industry value. We firstly do the unit root testing of air passenger volume and tertiary industry value.

Table 2 Unit root test for air passenger volume and tertiary industry value

Variables	Test of ADF	Type of test	critical value	Conclusion	Akaike info criterion
X	-1.806235	（c，t，2）	-3.3588**	Unstable	16.26194
ΔX	-1.005175	（c，0，2）	-2.7180**	Unstable	16.63612
$\Delta^2 X$	-4.542417	（0，0，0）	-2.7760	Stable	16.23886
Y	-1.609207	（c，t，3）	-3.3820**	Unstable	12.43998
ΔY	-3.457094	（c，t，1）	-3.8288*	Unstable	12.81008
$\Delta^2 Y$	-4.136648	（0，0，0）	-2.7760	stable	13.26774

(where * denotes significance in 5% level, ** denotes significance in 10% level, the remainder denote significance in 1% level. Minimum AIC denotes choosing lag order by AIC information criteria)

From Table 2, we could conclude that the optimal lag order of both air passenger volume and tertiary industry value is I (2). We use the Granger causality testing to check the relationship between air passenger volume and tertiary industry value. In general, the following VAR model is chosen:

$$X_t = a_0 + a_1 X_{t-1} + a_2 X_{t-2} + \cdots + a_k X_{t-k} + b_1 Y_{t-1} + b_2 Y_{t-2} + \cdots + b_k Y_{t-k} + e_{1t} \quad (1)$$

$$Y_t = c_0 + c_1 X_{t-1} + c_2 X_{t-2} + \cdots + c_k X_{t-k} + d_1 Y_{t-1} + d_2 Y_{t-2} + \cdots + d_k Y_{t-k} + e_{2t} \quad (2)$$

Here, k is lag order, while e_{1t} and e_{2t} are residuals respectively.

In order to select the right lag order, k, of the VAR model, we need to justify the independence and normality of the residual for each lag order model. We use the rule from generalization to specification to select the lag order, starting from high order and comparing the normality of residual for each order, and then select the appropriate lag order among them. After repeated computation, we will perform the Granger causality test at lag order 1.

Table 3 The result of Granger causality test

Lags: 1			
Null Hypothesis:	Obs	F-Statistic	Probability
X does not Granger Cause Y	15	0.24762	0.62775
Y does not Granger Cause X		6.24131	0.02800

Since the p-value is 0.02800, tertiary industry value has significant influence on air passenger volume at the 5% level; whereas air passenger volume doesn't have significant influence on tertiary industry value in 5% level due to $p = 0.62775$.

We diagnose these residuals and have the following results, shown as table 4.

Table 4 Independent and normality test form of the residuals

Residual	Jarque-Bera Probability （critical value is 5.99）	Probability of normal distribution	Result of normality test（5%）	Q value of independent	Probability of independent	Result of independent test（5%）
Residual of Y Equation	0.4805	0.7891	Normality	1.7928	0.181	independent
Residual of X Equation	4.1965	0.12268	Normality	0.5463	0.460	independent

From Table 4, the above residuals all satisfy independence and normality. It is reasonable to select the lag orders.

The equations of causality test are below:

$$X_t = -1072.619 + 0.353972 X_{t-1} + 0.728153 Y_{t-1} \quad (3)$$

[-1.752236] [1.116168] [2.498261]

(0.1052) (0.2862) (0.0280)

$$R^2 = 0.95, \quad F = 125.8999, \quad DW = 1.8$$

$$Y_t = 78.97601 - 0.055419 X_{t-1} + 1.117504 Y_{t-1} \qquad (4)$$

$$[0.367379] \quad [-0.497610] \quad [10.91784]$$

$$(0.7197) \quad (0.6278) \quad (0.0000)$$

$$R^2 = 0.99, \quad F = 1060.9, \quad DW = 1.24447$$

From the above, we can see that there is no inevitable correlation between the current air passenger volume and that with 1 lag, while it will be significantly affected by the tertiary industry value with lag 1. In other words, when the tertiary industry value increases by 1%, the next period's volume will increase by 0.73%. This kind of significance indicates that the scale of the urban circulation and services have great impacts on air flights scheduling.

4. Conclusion

CCAN is a small-world network, where its degree distribution follows a double power-law and there is a clear hierarchical layout. In CCAN's development process, there are a variety of external factors affecting the connections. This paper analyzes the relationship between air passenger traffic and three kinds of external factors such as population, economic size and distance. We find that the tertiary industry has a very important influence on the Chinese air passenger traffic. When the tertiary industry value increases by 1%, the next period's volume will increase by 0.73%. The population exhibits a weakly positive correlation with the number of flights, yet it did not show obvious correlation with transportation capacity. In CCAN, the flight distances assemble in the range from 400km to 1400km. When the distance exceeds 1400km, the airline number decays very fast with the increasing of distance. In a certain distance range (<2000km), there is no clear correlation between distance and flow. It is worth emphasizing that, as sudden acceleration construction of high-speed railway and passenger dedicated line (PDL), the current pattern of China's transportation industrial pattern has profoundly changed. The short-haul routes are being serious challenged by high-speed railway and PDL. The impact of distance on airline network will present a new features, which is one of our future research interests.

Reference

[1] S. J. Hu, Transportation Studies (China Communications Press, Beijing, 2001).
[2] A. R.Goetz, Growth and Change 23 (1992) 217-238.
[3] M. D. Irwin, J. D.Kasarda, American Sociological Review 56 (1991) 524-537.
[4] K. G. Debbage, J. Air Transp. Manag. 5 (1999) 211-221.
[5] E. J. Taaffe, Econ. Geogr. 38 (1962) 1-14.
[6] M. W. Horner, M. E. O'Kelly, J. Transp. Geogr. 9 (2001) 255-265.
[7] J. Bowen, J. Transp. Geogr. 8 (2000) 25-41.
[8] K. O'Connor, J. Transp. Geogr. 3 (1995) 269-279.
[9] K. Button, S. Lall, R. Stough, M. Trice, J. Air Transp. Manag. 5 (1999) 53-59.
[10] O. W. Wojahn, Transport. Res. E 37 (2001) 267-279.
[11] G. Nero, Transport. Res. E 35 (1999) 225-239.
[12] M. E. O'Kelly, Eur. J. Oper. Res. 32 (1987) 393-404.
[13] M. E. O'Kelly, J. Transp. Geogr. 6 (1998) 171-186.
[14] T. Aykin, Eur. J. Oper. Res. 79 (1994) 501-523.
[15] T. Aykin, Eur. J. Oper. Res. 83 (1995) 200-219.
[16] L. A. N. Amaral, A. Scala, M. Barthélemy, H. E. Stanley, Proc. Natl. Acad. Sci. 97 (2000) 11149-11152.

[17] R. Guimerà, L. A. N. Amaral, Eur. Phys. J. B 38 (2004) 381-385.
[18] A. Barrat, M. Barthélemy, A. Vespignani, J. Stat. Mech. 5 (2005) P05003.
[19] H. K. Liu, T. Zhou, Prog. Nat. Sci. 18 (2008) 601-608.
[20] H. K. Liu, T. Zhou, Acta Phys. Sin. 56 (2007) 106-112.
[21] G. K. Zipf, Selective Studies and the Principle of Relative Frequency in Language (Harvard University Press, Cambridge, 1932).
[22] H. K. Liu, X. L. Zhang, L. Cao, B. H. Wang, T. Zhou, Sci. China Ser. G 39 (2009) 935-942.
[23] G. Bianconi, P. Pinb, M. Marsili, Proc. Natl. Acad. Sci. 106 (2009) 11433-11438.
[24] S-H. Yook, H. Jeong, A-L. Barabási, Proc. Natl. Acad. Sci. 99 (2002) 13382-13386.
[25] A. F. Rozenfeld, R. Cohen, D. ben-Avraham, S. Havlin, Phys. Rev. Lett. 89 (2002) 218701.
[26] J. Jost, M. P. Joy, Phys. Rev. E 66 (2002) 036126.
[27] J. Dall, M. Christensen, Phys. Rev. E 66 (2002) 016121.
[28] P. Sen, K. Banerjee, T. Biswas, Phys. Rev. E 66 (2002) 037102.
[29] S. S. Manna, P. Sen, Phys. Rev. E 66 (2002) 066114.
[30] R. Xulvi-Brunet, I. M. Sokolov, Phys. Rev. E 66 (2002) 026118.
[31] M. Barthélemy, Europhys. Lett. 63 (2003) 915-921.
[32] D. ben-Avraham, A. F. Rozenfeld, R. Cohen, S. Havlin, Physica A 330 (2003) 107-116.
[33] S. Porta, P. Crucitti, V. Latora, Physica A 369 (2006) 853-866.
[34] M. T. Gastner, M. E. J. Newman, Eur. Phys. J. B 49 (2006) 247-252.
[35] M. Rosvall, A. Trusina, P. Minnhagen, K. Sneppen, Phys. Rev. Lett. 94 (2005) 028701.
[36] A. Cardillo, S. Scellato, V. Latora, S. Porta, Phys. Rev. E 73 (2006) 066107.
[37] V. Latora, M. Marchiori. Physica A 314 (2002) 109-113.
[38] P. Sen, S. Dasgupta, A. Chatterjee, P. A. Sreeram, G. Mukherjee, S. S. Manna, Phys. Rev. E 67 (2003) 036106.
[39] M. Kurant, P. Thiran, Phys. Rev. Lett. 96 (2006) 138701.
[40] M. Kurant, P. Thiran, Phys. Rev. E 74 (2006) 036114.
[41] D. Brockmann, L. Hufnage, T. Geisel, Nature 439 (2006) 462-465.